HIERARCHICAL BAYESIAN INFERENCE OF COSMIC SHEAR AND GALAXY SIZE MORPHOLOGY RELATION

OR INFERRING COVARIANCE WHEN YOUR OBSERVATIONS ARE INDIRECT AND PARAMETERS OF INTEREST MIGHT FOLLOW DIFFERENT DISTRIBUTIONS

Karan Shah
kshah84
CSE 8803 IUQ Final Report
Spring 2019

May 3, 2019

ABSTRACT

The gravitational lensing of distant galaxies by mass along the line-of-sight (cosmic shear) is a sensitive probe of both the expansion and structure growth rates of the Universe. The massive amount of data from the Large Synoptic Survey Telescope (LSST) and other surveys can be used to infer the properties of cosmic shear and the intrinsic properties of the distorted galaxy populations. We describe how Hierarchical Bayesian Models can be used to infer these properties while marginalizing nuisance parameters.

1 Introduction

The presence of mass (dark and luminous matter) in the line-of-sight path of light distorts the observed images of light sources. The distribution of this foreground mass can be inferred by measuring slight correlations in the observed properties of light sources in a given patch of the sky. Estimating this shear is not trivial as we do not know the intrinsic properties of the light sources.

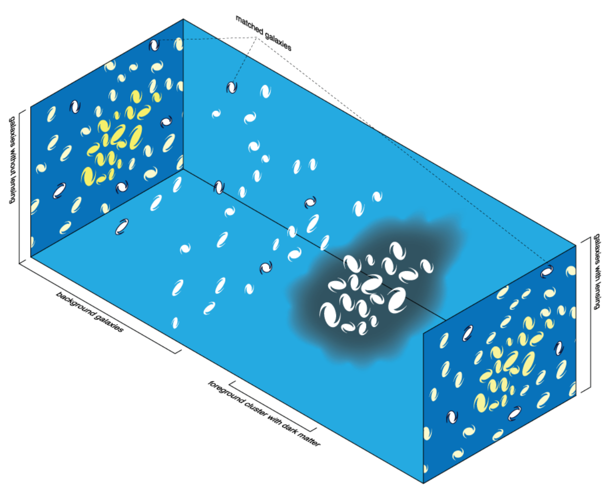

Figure 1: Geometry for gravitational lensing of distant galaxies.

This distortion caused by mass in the path of light is of the order of $\sim 0.1\%$ the instrinsic shape of galaxies. If the intrinsic shape of a galaxy is denoted by ϵ_{int}, the distorted ellipticity ϵ_{sh} observed under the presence of cosmic shear g is given by

$$\epsilon_{sh} \approx \epsilon_{int} + \epsilon_{sh} \tag{1}$$

Fig 2. shows an exagerrated version of this gravitaitonal lensing.

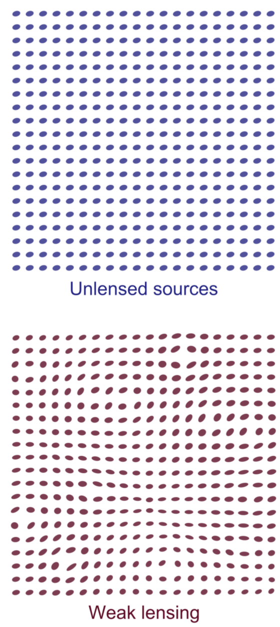

Figure 2: Intrinsic Ellipticities (blue) and Sheared Ellipticities (purple)

Inferring the cosmic shear from observed data is an intractable inverse modeling problem. By using a hierarchical bayesian model, we can take effects due to different factors like image noise, atmospheric distortions and cosmic shear into account infer mass distribution and intrinsic properties. The model also allows us to update our beliefs with incoming data and to marginalize out unknowable nuisance parameters. The forward process is shown in Fig 2.

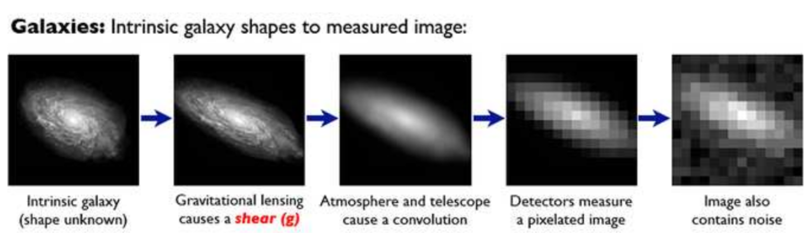

Figure 3: Illustration of the forward process.

Our observations are "postage stamps" of galaxy images. Each postage stamp contains 10-100 galaxies. For this report, we assume our observations are observed ellipcities ϵ_{obs} and other galaxy properties like flux (ϕ), redshift (z) and radius (r), infered from galaxy images with a noise parameter Σ_n that accounts for the last three stages of the forward process. Our goal is to determine the cosmic shear g and infer the covariances between the intrinsic galaxy properties (eg bigger galaxies might be brighter, but they might also be farther).

While we are conveniently pushing the last three stages of the forward process into Σ_n, an extended treatment of this problem can be found in [1].

2 Statistical Framework

2.0.1 Univariate Toy Model

We begin with a simple spherical cow toy problem. Given observations D_i, which is the observed ellipticity for galaxy i, we want to infer shear g. The intrinsic ellipcities are normally distributed with variance σ_e^2. The generating process in this case is defined as:

$$P(D|\epsilon_{sh}, \sigma_n) = N_D(\epsilon_{sh}, \sigma_n^2)$$
$$P(\epsilon_{int}|\sigma_e^2) = N_{\epsilon_{int}}(0, \sigma_e^2)$$
$$P(\epsilon_{int}(\epsilon_{sh}, g)|\sigma_e^2) = N_{(\epsilon_{sh}-g)}(0, \sigma_e^2)$$

with $\epsilon_{sh} \approx \epsilon_{int} + g$ whereσ_n^2 is the noise induced in ϵ_{sh} due to factors such as atmospheric distortion and we want to infer σ_e^2 and g from observations of N_gal galaxies.

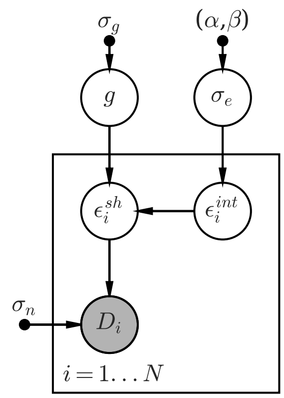

Figure 4: Probabilistic Graphical Model of our toy model.

For a single galaxy the likelihood is given by,

$$P(D|\epsilon_{int}, g) = P(D|\epsilon_{sh}(\epsilon_{int}, g)) \tag{2}$$

We don't need to know the unsheared shapes of each galaxy. Marginalizing ϵ_{int} out,

$$P(D|g) = \int_{-\infty}^{\infty} d\epsilon_{int} Pr(D|\epsilon_{int}, g, \sigma_n) Pr(\epsilon_{int}|I) \tag{3}$$

$$P(D|g) = \frac{e^{-\frac{(D_i-g)^2}{2(\sigma_e+\sigma_n)^2}}}{\sqrt{2\pi\left(\sigma_e+\sigma_n\right)^2}} \tag{4}$$

For N_{gal} galaxies,

$$P(\{D_i\}_1^{N_{gal}}|g) = \prod_{i=1}^{N_{gal}} \frac{e^{-\frac{(D_i-g)^2}{2(\sigma_e+\sigma_n)^2}}}{\sqrt{2\pi\left(\sigma_e+\sigma_n\right)^2}} \tag{5}$$

which can be written as

$$\left(\frac{e^{-\frac{(-g)^2}{2(\sigma_e+\sigma_n)^2}}}{\sqrt{2\pi\left(\sigma_e+\sigma_n\right)^2}}\right)^{N_{gal}} * \prod_{i=1}^{N_{gal}} e^{(2D_ig-D_i^2)} \tag{6}$$

Simplifying,

$$P(\{D_i\}_1^{N_{gal}}|g) = (N_g(0, (\sigma_e+\sigma_n)^2))^{N_{gal}} * e^{-(\overline{D}^2-2gN_{gal}\overline{D}+\sigma_n^2)} \tag{7}$$

where $\overline{D} = \overline{\epsilon}_{sh}$ is the mean of D distribution

Using this likelihood and appropriate conjugate priors like the following:

$$\pi(g) = N_g(0, \sigma_g^2)$$

$$\pi(\sigma_e^2) = \Gamma^{-1}_{sigma_e^2}(a, b)$$

where Γ^{-1} is the Inv. Gamma distribution

We can easily use Gibbs sampling to estimate g and σ_e^2 with rapid convergence.

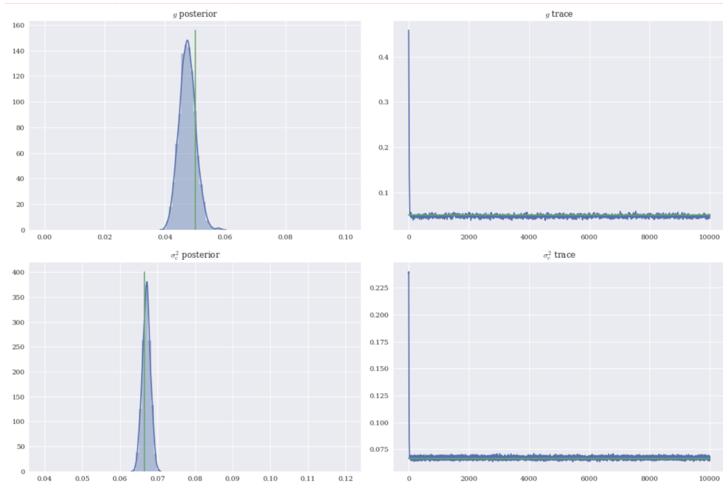

Figure 5: Gibbs Sampling results from the toy model. Green line denotes true value.

2.0.2 Hairier Model

The simple toy model works well when there is plentiful of data and we are only observing one ellipticity ϵ_{obs}. In actual observations, there are two shape parameters ϵ_1 and ϵ_2 that are distorted by g and other non-shape parameters like flux, radius etc which are not distorted by g and might not follow gaussian distribution. Moreover, galaxy postage stamps often have very few ($\approx$ 10 data points. We use importance sampling with an interim prior to augment our observations.

Consider this problem: Our galaxy image model parameters are, $\omega = [\epsilon_1, \epsilon_2, \nu, r, \phi]$, where

$$\epsilon_1 \sim N(0, \sigma_{e1}).$$
$$\epsilon_2 \sim N(0, \sigma_{e2}).$$
$$\nu \sim N(\mu_\nu, \sigma_\nu).$$
$$r \sim lognorm(\mu_r, \sigma_r).$$
$$\phi \sim powerlaw(a).$$

In the weak lensing regime, $\epsilon_{sh} \approx \epsilon_{int} + g$.

$$\omega_{sh} = [\epsilon_1 + g, \epsilon_2 + g, \nu, r, \phi].$$

We can transform the non-Gaussian distributions into approximately Gaussian distributions via,

$$g(r) = \log(r).$$

$$h(\phi) = \log(\frac{\phi^a}{1 - \phi^a}).$$

So the transformed data is,

$$\omega' \equiv [\epsilon_1, \epsilon_2, \nu, g(r), h(\phi)].$$
$$\omega'_{sh} = [\epsilon_1 + g, \epsilon_2 + g, \nu, g(r), h(\phi)].$$

where $\omega' \sim N(\alpha, \Sigma_{int})$. We define this transformation as

$$\omega' = f(\omega).$$

with Jacobian J. Then,

$$J \equiv \left|\frac{d\omega'}{d\omega}\right| = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} & 0 \\ 0 & 0 & 0 & 0 & \frac{a}{\phi(1-\phi^a)} \end{vmatrix} \tag{8}$$

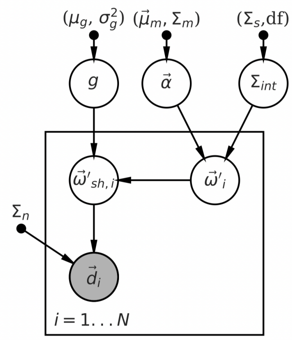

Figure 6: PGM for transformed parameters

Importance Sampling

The monte carlo integration for some integral $\int f(x)p(x)dx$ is:

$$\int f(x)p(x)dx \approx \frac{1}{N}\sum_{i=1}^{N} f(x_i), x_i \sim p(x). \tag{9}$$

But sometimes we can't sample from $p(x)$ or don't get sufficiently precise results with it. In that case, we use importance sampling, by using a distribution $q(x)$ that is easier to sample from.

$$\int f(x)p(x)dx = \int \frac{f(x)p(x)}{q(x)} q(x)dx \approx \frac{1}{N}\sum_{i=1}^{N} \frac{p(x_i)}{q(x_i)} f(x_i), x_i \sim q(x). \tag{10}$$

We added weights $\frac{p(x_i)}{q(x_i)}$ to the monte carlo sum in Eqn 1.

Approximating the Likelihood

For inference of galaxy distribution parameters, we aim to evaluate the marginal likelihood,

$$P(\mathbf{D}|\alpha, \Sigma_e, g) \propto \int d\omega_i P(\omega_i(\omega_{sh}, g)|\alpha, \Sigma_e, g) P(\mathbf{D}|\omega). \tag{11}$$

Transforming ω to ω',

$$P(\mathbf{D}|\alpha, \Sigma_e, g) \propto \int d\omega' \left|\frac{d\omega}{d\omega'}\right| P(\omega'(\omega'_{sh}, g)|\alpha, \Sigma_e, g) P(\mathbf{D}|\omega'). \tag{12}$$

Using an interim prior I,

$$P(\mathbf{D}|\Sigma) = \int d\omega' \left|\frac{d\omega}{d\omega'}\right| \left[P(\mathbf{D}|\omega_{sh} = f^{-1}(\omega'_{sh}))P(\omega_{sh}|I)\right] \frac{P(\omega'(\omega'_{sh}, g)|\alpha, \Sigma_{int}, g)}{P(\omega_{sh}|I)}., \tag{13}$$

we get the likelihood for all galaxies via importance sampling as follows,

$$P(\{\mathbf{D}_i\}_{i=1}^N|\Sigma) \propto \prod_{i=1}^N \frac{Z_i}{K} \sum_{j=1}^K J^{-1}(\omega_{ij}) \frac{P(\omega'_{ij}(\omega'_{sh} = f(\omega_{sh}), g)|\alpha, \Sigma_{int}, g)}{P(\omega_{ij,sh}|I)}. \tag{14}$$

where $\omega_{sh} \sim \left[P(\mathbf{d}|\omega_{sh} = f^{-1}(\omega'_{sh}))Pr\omega_{sh}|I)\right]$ are the K samples drawn.